

- REGRESSION ANALYSIS
- LOGNORMAL DISTRIBUTION

MEK 04/05

ALI
S

**SELANG KEPERCAYAAN KURVA REGRESI
NONPARAMETRIK DENGAN ERROR LOGNORMAL
BERDASARKAN ESTIMATOR KERNEL**

SKRIPSI

**MILIK
PERPUSTAKAAN
UNIVERSITAS AIRLANGGA
SURABAYA**



KAMELIA ALI

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS AIRLANGGA
SURABAYA
2004**

**SELANG KEPERCAYAAN KURVA REGRESI
NONPARAMETRIK DENGAN ERROR LOGNORMAL
BERDASARKAN ESTIMATOR KERNEL**

SKRIPSI

**Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Sains
Bidang Matematika di Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Airlangga**

Oleh :

KAMELIA ALI
NIM. 080012079

Tanggal Lulus : 24 Desember 2004

Disetujui Oleh :

Pembimbing I



Nur Chamidah, S.Si, M.Si.
NIP. 132 205 653

Pembimbing II



Drs. Ardi Kurniawan, M.Si.
NIP. 132 230 977

LEMBAR PENGESAHAN SKRIPSI

Judul : SELANG KEPERCAYAAN KURVA REGRESI
NONPARAMETRIK DENGAN ERROR LOGNORMAL
BERDASARKAN ESTIMATOR KERNEL

Penyusun : KAMELIA ALI

NIM : 080012079

Tanggal Ujian : 24 DESEMBER 2004

Disetujui Oleh :

Pembimbing I



Nur Chamidah, S.Si, M.Si.
NIP. 132 205 653

Pembimbing II

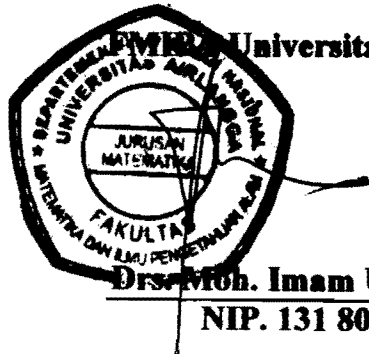


Drs. Ardi Kurniawan, M.Si.
NIP. 132 230 977

Mengetahui :

Ketua Jurusan Matematika

FMIPA Universitas Airlangga



Drs. Mch. Imam Utoyo, M.Si
NIP. 131 801 397

Kamelia Ali, 2004. *Selang Kepercayaan Kurva Regresi Nonparametrik dengan Error Lognormal Berdasarkan Estimator Kernel*. Skripsi ini dibawah bimbingan Nur Chamidah S.Si, M.Si dan Drs. Ardi Kurniawan M.Si, Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Airlangga.

ABSTRAK

Analisis regresi adalah salah satu alat di dalam statistika yang umumnya digunakan untuk menentukan hubungan antara dua variabel yaitu variabel respon Z dan variabel prediktor X untuk n pengamatan. Hubungan variabel tersebut dapat dinyatakan sebagai berikut :

$$Z_i = g(x_i)e_i \quad i = 1, 2, \dots, n \quad e_i \sim A(0, \sigma^2)$$

dengan $g(x_i)$ kurva regresi yang tidak diketahui. Model tersebut akan ditransformasi dengan cara dilogaritmakan sehingga dihasilkan model

$$Y_i = m(x_i) + \varepsilon_i \quad i = 1, 2, \dots, n \quad \varepsilon_i \sim N(0, \sigma^2)$$

Tujuan dari penulisan ini adalah untuk mengkonstruksi selang kepercayaan kurva regresi dengan menggunakan pendekatan kernel. Secara umum selang kepercayaan dari kurva regresi dengan pendekatan estimator kernel adalah sebagai berikut :

$$P \left[\frac{\exp\left(\frac{\frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) y_i}{\frac{1}{n} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)}\right)}{\exp\left[Z_{\alpha/2} \left(\frac{\sigma^2}{\left[\frac{1}{n} \sum_{i=1}^n K_h(x-x_i)\right] nh}\right)^{1/2}}\right]} < g(x_i) < \frac{\exp\left(\frac{\frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) y_i}{\frac{1}{n} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)}\right)}{\exp\left[Z_{\alpha/2} \left(\frac{\sigma^2}{\left[\frac{1}{n} \sum_{i=1}^n K_h(x-x_i)\right] nh}\right)^{1/2}}\right]} = 1 - \alpha \right]$$

Dari penerapan model regresi nonparametrik pada data pohon *Gmelina Arborea Roxb*, untuk hubungan antara volume sebagai variabel respon dan diameter sebagai variabel prediktor diperoleh nilai estimator, nilai batas atas dan batas bawah selang kepercayaannya seperti dalam lampiran 5 dengan h optimal berdasarkan kriteria GCV adalah 0.32 dan gambar selang kepercayaannya ditunjukkan oleh gambar 2. Sedangkan untuk hubungan antara volume sebagai variabel respon dan tinggi sebagai variabel prediktor diperoleh nilai estimator, nilai batas atas dan batas bawah selang kepercayaannya seperti dalam lampiran 6 dengan h optimal berdasarkan kriteria GCV adalah 0.3 dan gambar selang kepercayaannya ditunjukkan oleh gambar 4.

Kata Kunci: Selang Kepercayaan, Regresi Nonparametrik, estimator Kernel, Distribusi Lognormal.

Kamelia Ali, 2004. *Confidence Interval of Nonparametric Regression Curve with Lognormal Error based on Kernel Estimator Approach*. This *skripsi* under guidance of Nur Chamidah S.Si, M.Si and Drs. Ardi Kurniawan M.Si, Department of Mathematics. Faculty of Mathematics and Natural Science Airlangga University

ABSTRACT

Regression analysis is one of statistical which usually used to determine relationship between two variables i.e response variables Z and predictor variables X for n observation. The relationship of this variables can be modeled as :

$$Z_i = g(x_i)e_i \quad i = 1, 2, \dots, n \quad e_i \sim N(0, \sigma^2)$$

Where $g(x)$ is an unknown regression function. The model would be transformed by taking natural logarithm. Such that, we obtain

$$Y_i = m(x_i) + \varepsilon_i \quad i = 1, 2, \dots, n \quad \varepsilon_i \sim N(0, \sigma^2)$$

The purpose of this *skripsi* is to construct confidence interval of nonparametric regression curve with lognormal errors based on kernel estimator approach. The confidence interval of nonparametric regression curve with lognormal error based on kernel estimator approach is:

$$P \left(\frac{\exp \left(\frac{\frac{1}{nh} \sum_{i=1}^n K \left(\frac{x-x_i}{h} \right) y_i}{\frac{1}{n} \sum_{i=1}^n K \left(\frac{x-x_i}{h} \right)} \right)}{\exp \left(Z_{\alpha/2} \left(\frac{\sigma^2}{\left[\frac{1}{n} \sum_{i=1}^n K_s(x-x_i) \right] nh} \right)^{1/2} \right)} < g(x_i) < \exp \left(\frac{\frac{1}{nh} \sum_{i=1}^n K \left(\frac{x-x_i}{h} \right) y_i}{\frac{1}{n} \sum_{i=1}^n K \left(\frac{x-x_i}{h} \right)} \right) \exp \left(Z_{\alpha/2} \left(\frac{\sigma^2}{\left[\frac{1}{n} \sum_{i=1}^n K_s(x-x_i) \right] nh} \right)^{1/2} \right) \right] = 1 - \alpha$$

The nonparametric regression model would be applied to *Gmelina Arborea Roxb tree* data with volume as response variable and diameter as predictor variables. Hence, we obtained the estimator, upper and lower values of confidence interval as given in appendix 5. Optimum bandwidth based on GCV criterion is 0.32 and the confidence interval is shown in figure 2. And for volume as response variable and diameter as predictor variables. Hence, we obtained the estimators, upper and lower values of confidence interval as given in appendix 6. Optimum bandwidth based on GCV criterion is 0.32 and the confidence interval is shown in figure 4.

Key Words: Confidence Interval, Nonparametric Regression, Kernel Estimator, Lognormal Distribution.